

# On the Exact BER of Bit-Wise Demodulators for One-Dimensional Constellations

Mikhail Ivanov, Fredrik Brännström, Alex Alvarado, Erik Agrell

**Abstract**—The optimal bit-wise demodulator for  $M$ -ary pulse amplitude modulation (PAM) over the additive white Gaussian noise channel is analyzed in terms of uncoded bit-error rate (BER). New closed-form BER expressions for 4-PAM with any labeling are developed. Moreover, closed-form BER expressions for 11 out of 23 possible bit patterns for 8-PAM are presented, which enable us to obtain the BER for 8-PAM with some of the most popular labelings, including the binary reflected Gray code and the natural binary code. Numerical results show that, regardless of the labeling, there is no difference between the optimal demodulator and the symbol-wise demodulator for any BER of practical interest (below 0.1).

**Index Terms**—Additive white Gaussian noise channel, binary reflected Gray code, bit error probability, bit-interleaved coded modulation, demapper, demodulator, LLRs, logarithmic likelihood ratio, pulse-amplitude modulation, uncoded transmission.

## I. INTRODUCTION AND MOTIVATION

Current wireless communication systems are based on the bit-interleaved coded modulation (BICM) paradigm introduced in [1] and later studied in [2], [3]. One key element in these systems is the demodulator which calculates logarithmic likelihood ratios (LLR, also known as L-values) for the received bits, which are then passed to the channel decoder. The calculation of L-values is crucial in many other coded systems. The coded performance analysis of BICM systems is generally not straightforward and is usually carried out either numerically by Monte-Carlo simulation or in terms of lower and upper bounds [2, Sec. 4], [3, Ch. 4]. In this paper, we analyze the *uncoded* performance of bit-wise demodulators over the additive white Gaussian noise (AWGN) channel.

The optimal bit-wise demodulator (BD) minimizing the BER implies the calculation of (exact) L-values for the received bits. The uncoded performance of such a demodulator has been studied in [4], where closed-form expressions for the BER for 4-PAM with the binary reflected Gray code (BRGC) [5]–[7] are presented. Due to the complexity of the BD, the calculation of L-values in practical systems is usually done based on the so-called max-log approximation [8, eq. (5)], [9, eq. (1)]. We call this demodulator the approximate BD (ABD). The ABD is equivalent to the symbol detector in terms

of uncoded BER [10, Sec. IV-A], whose performance is well documented in literature, e.g., [11, Ch. 5], [12, Ch. 10], [6], [13]–[18] and references therein.

It is well known that the uncoded BER of one-dimensional constellation can be expressed as a sum of Gaussian Q-functions, cf. [11, Ch. 5], [12, Ch. 10] and references therein. The arguments of the Q-functions depend on the points that separate the decision regions associated to different bits. We refer to these points as thresholds. The computation of the thresholds for the BD—the optimal bit-wise demodulator—is in general complicated and unknown. In this paper, however, we show that this problem can be solved analytically for 4-PAM and any labeling, extending the results presented in [4]. Moreover, we also analytically calculate the thresholds for 8-PAM with some relevant labelings, including the BRGC, the natural binary labeling (NBC) [19, Sec. II-B], the folded binary code (FBC) [16] [19, Sec. II-B], the binary semi-Gray code (BSGC) [19, Sec. II-B], and the anti-Gray code (AGC) [20]. Numerical results show that optimal and suboptimal demodulators are different in terms of the BER only at a very low SNR. At BER below 0.1 there is no notable difference between them.

The rest of the paper is organized as follows. In Sec. II we introduce the notation convention, the system model, and the two demodulators. In Sec. III the BER analysis is presented. The patterns that form a labeling are studied in Sec. IV. The threshold computation for the BD is shown in Sec. V and the numerical results in Sec. VI. The conclusions are drawn in Sec. VII.

## II. PRELIMINARIES

### A. Notation Convention

In this paper the following notation is used. Lowercase letters  $x$  denote real or complex scalars and boldface letters  $\mathbf{x}$  denote a row vector of scalars. The complex conjugate of  $x$  is denoted by  $x^*$ . Blackboard bold letters  $\mathbb{X}$  denote matrices with elements  $x_{i,j}$  in the  $i$ th row and the  $j$ th column and  $(\cdot)^T$  denotes transposition. Calligraphic capital letters  $\mathcal{X}$  denote sets, where the set of real numbers is denoted by  $\mathcal{R}$ . The binary complement of  $x \in \{0, 1\}$  is denoted by  $\bar{x} = 1 - x$  and its bipolar representation by  $\tilde{x} = 2x - 1$ . Binary addition (exclusive-OR) of two bits  $a$  and  $b$  is denoted by  $a \oplus b$ . Random variables are denoted by capital letters  $X$  and probabilities by  $\Pr\{\cdot\}$ . The Gaussian Q-function is defined as  $Q(x) \triangleq (1/\sqrt{2\pi}) \int_x^\infty \exp(-t^2/2) dt$ .

Research supported by the Swedish Research Council, Sweden (under grant #621-2006-4872 and #621-2011-5950) and by the European Community's Seventh's Framework Programme (FP7/2007-2013) under grant agreement No. 271986.

M. Ivanov, F. Brännström, and E. Agrell are with the Dept. of Signals and Systems, Chalmers Univ. of Technology, SE-41296 Göteborg, Sweden (emails: {mikhail.ivanov,fredrik.brannstrom,agrell}@chalmers.se). A. Alvarado is with the Dept. of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom (email: alex.alvarado@ieee.org).

### B. System Model

In this paper we analyze a system where a vector of binary data  $\mathbf{b} = [b_1, \dots, b_m]$  is fed to a modulator. The modulator carries out a one-to-one mapping from  $\mathbf{b}$  to one of the  $M$  constellation points  $x \in \mathcal{X} = \{s_1, \dots, s_M\}$  for transmission over the physical channel, where  $M = 2^m$ . We assume that  $s_1 < s_2 < \dots < s_M$ . The modulator is defined as the function  $\Phi : \{0, 1\}^m \rightarrow \mathcal{X}$ .

The modulator is defined by the constellation and its binary labeling. A binary labeling is specified by the matrix  $\mathbb{C} = [\mathbf{c}_1^T, \dots, \mathbf{c}_M^T]^T$  of dimensions  $M$  by  $m$ , where the  $i$ th row  $\mathbf{c}_i = [c_{i,1}, \dots, c_{i,m}]$  is the binary label of the constellation point  $s_i$ , i.e.,  $\Phi(\mathbf{c}_i) = s_i$ .

For PAM constellations,  $s_i = -d(M - 2i + 1)$ ,  $i = 1, \dots, M$ , where  $d = \sqrt{3/(M^2 - 1)}$  to normalize the constellation to unit average energy, i.e.,  $E_s = (1/M) \sum_{i=1}^M s_i^2 = 1$ . We assume the bits to be independent and identically distributed (i.i.d.) with  $\Pr\{B_j = u\} = 0.5, \forall j$  and  $u \in \{0, 1\}$ , and thus, the symbols are equiprobable, i.e.,  $\Pr\{X = s_i\} = 1/M, \forall i$ .

In this paper we consider a discrete time memoryless AWGN channel with output  $y = x + \eta$ , where  $x \in \mathcal{X}$  and the noise sample  $\eta$  is a zero-mean Gaussian random variable with variance  $N_0/2$ . The conditional probability density function (PDF) of the channel output given channel input is

$$p_{Y|X}(y|x) = \sqrt{\frac{\gamma}{\pi}} e^{-\gamma(y-x)^2}, \quad (1)$$

where the average signal to noise ratio (SNR) is defined as  $\gamma \triangleq E_s/N_0 = 1/N_0$ .

The observation  $y$  is used by the demodulator to decide on the received binary sequence, i.e., to produce  $\hat{\mathbf{b}} = [\hat{b}_1, \dots, \hat{b}_m]$ . In this paper we consider two demodulators to obtain  $\hat{\mathbf{b}}$  from  $y$ , which are described in the next section.

### C. Demodulators

The BD calculates (a posteriori) L-values for the  $m$  bits based on the observation  $y$ , i.e.,

$$l_j(y) \triangleq \log \frac{\Pr\{B_j = 1|Y = y\}}{\Pr\{B_j = 0|Y = y\}} \quad (2)$$

$$= \log \frac{\sum_{x \in \mathcal{X}_{j,1}} e^{-\gamma(y-x)^2}}{\sum_{x \in \mathcal{X}_{j,0}} e^{-\gamma(y-x)^2}}, \quad (3)$$

for  $j = 1, \dots, m$  and  $\mathcal{X}_{j,u} \triangleq \{s_i \in \mathcal{X} : c_{i,j} = u, \forall i\}$ . To pass from (2) to (3) Bayes' rule was used together with the i.i.d. assumption of the bits and the conditional PDF in (1). The BD uses the L-values in (3) to make a decision on the received bit according to the rule

$$\hat{b}_j^{\text{BD}} = \begin{cases} 1 & \text{if } l_j(y) \geq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The implementation of the BD in its exact form (3) is complicated, especially for large constellations, as it requires calculation of the logarithm of a sum of exponentials. To overcome this problem, approximations are usually used in practice. The most common approximation is the so-called

max-log approximation ( $\log \sum_i e^{\lambda_i} \approx \max_i \lambda_i$ ) [1, eq. (3.2)], [2, eq. (9)], [8, eq. (5)], [21, eq. (8)], which used in (3) gives

$$\tilde{l}_j(y) = \gamma \left[ \min_{x \in \mathcal{X}_{j,0}} (y-x)^2 - \min_{x \in \mathcal{X}_{j,1}} (y-x)^2 \right]. \quad (5)$$

The use of the max-log approximation transforms the nonlinear relationship (3) into a piece-wise linear relationship (5), as previously shown in e.g., [22, Fig. 3], [23, eqs. (11)–(14)].

The ABD is defined as the demodulator that applies the same decision rule (4) based on L-values calculated by (5). As mentioned in [10, Sec. IV-A] the ABD is equivalent to the symbol detector in terms of uncoded BER.

### III. BER FOR ONE-DIMENSIONAL CONSTELLATIONS

The BER for a given labeling  $\mathbb{C}$  can be expressed as

$$P_{\mathbb{C}} = \frac{1}{m} \sum_{j=1}^m P_j, \quad (6)$$

where using the law of total probability, the BER for the  $j$ th bit position  $P_j \triangleq \Pr\{\hat{B}_j \neq b_j | B_j = b_j\}$  can be written as

$$P_j = \frac{1}{M} \sum_{i=1}^M \Pr\{\hat{B}_j \neq c_{i,j} | X = s_i\}. \quad (7)$$

The BER for the  $j$ th bit position  $P_j$  depends only on the subconstellations  $\mathcal{X}_{j,0}$  and  $\mathcal{X}_{j,1}$  (cf. (3)–(5)), i.e.,  $P_j$  is a function of the  $j$ th column of  $\mathbb{C}$ .

We define a bit pattern (or simply pattern) as a length- $M$  binary vector  $\mathbf{p} = [p_1, \dots, p_M] \in \{0, 1\}^M$  with Hamming weight  $M/2$ . The labeling  $\mathbb{C}$  can now be defined by  $m$  patterns, each corresponding to one column of  $\mathbb{C}$ . We index the patterns as  $\mathbf{p}_w$  with  $w$  being the decimal representation of the vector  $\mathbf{p}$ , i.e.,  $w = \sum_{i=1}^M 2^{M-i} p_i$ . For example, for  $M = 4$ , the pattern  $[0, 1, 0, 1]$  is indexed as  $\mathbf{p}_5$ . The BER for the labeling  $\mathbb{C}$  does not depend on the order of its columns, and thus, the BER for the labeling  $\mathbb{C}$  is fully determined by a set of  $m$  patterns (indices)  $\mathcal{W} = \{w_1, \dots, w_m\}$ . Based on the previous discussion, from now on we concentrate our analysis only on patterns (and not on labelings).

To analyze the BER of a pattern (PBER), the observation space  $\mathcal{R}$  is split into two disjoint decision regions, i.e.,  $\mathcal{Y}_0 = \{y \in \mathcal{R} : \hat{b} = 0\}$  and  $\mathcal{Y}_1 = \{y \in \mathcal{R} : \hat{b} = 1\}$  such that  $\mathcal{Y}_0 \cup \mathcal{Y}_1 = \mathcal{R}$ . Using the definition of  $\mathcal{Y}_0$  and  $\mathcal{Y}_1$ , the PBER for the pattern  $\mathbf{p} = [p_1, \dots, p_M]$  can be rewritten as

$$P = \frac{1}{M} \sum_{i=1}^M \Pr\{Y \in \mathcal{Y}_{\bar{p}_i} | X = s_i\}. \quad (8)$$

By expressing  $P$  as in (8), it is clear that the PBER in (7) can be calculated using the decision region  $\mathcal{Y}_0$  only, as opposed to alternative approaches where (8) is expressed in terms of the PDF of the L-values [24, eq. (19)].

#### A. Decision Thresholds

One key element in the BER analysis presented in this paper is the *decision thresholds*. Decision thresholds (or simply thresholds) for a given pattern  $\mathbf{p}$  are the points that separate the decision regions for zeros and ones, and thus, they determine

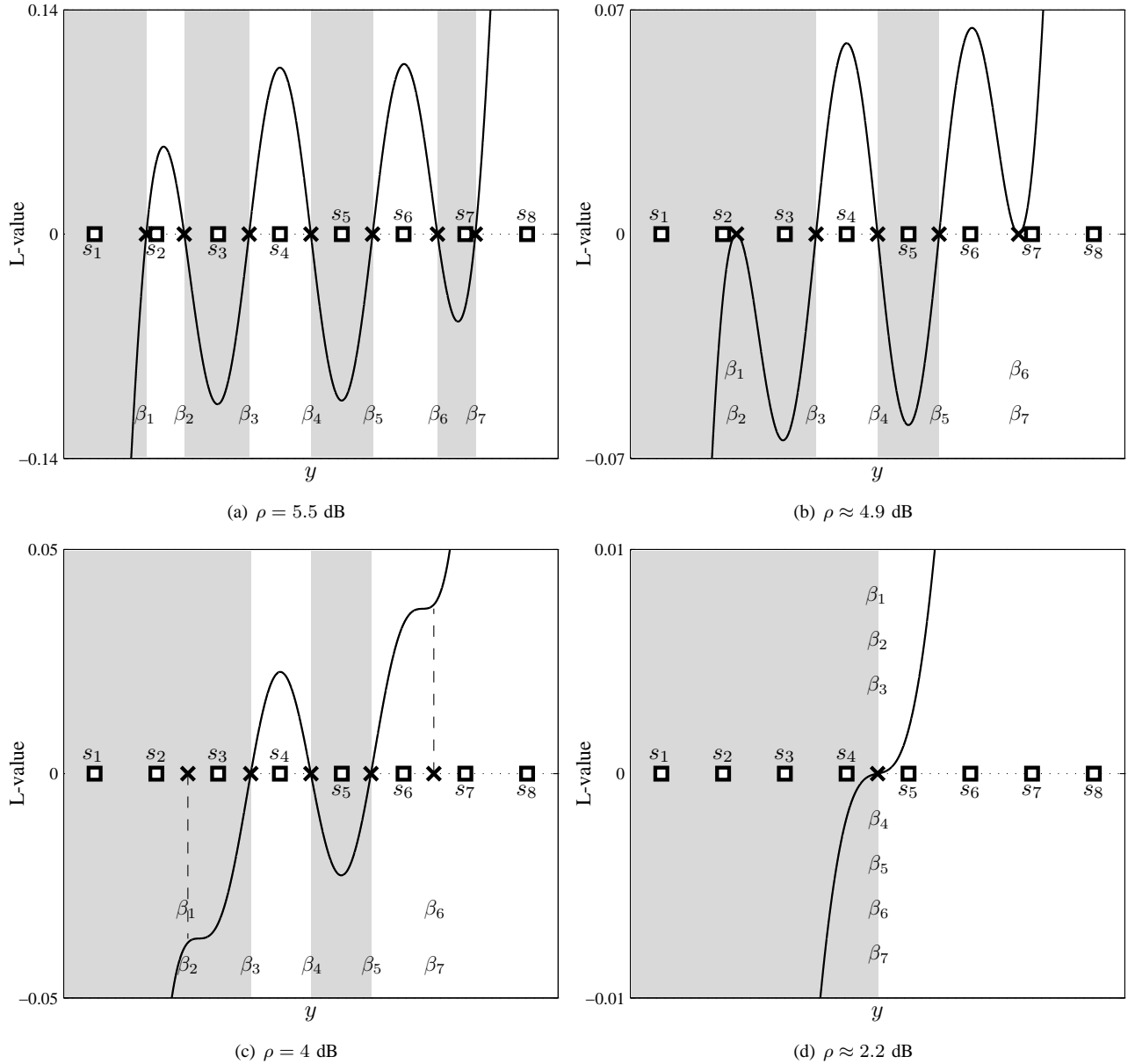


Fig. 1. L-values in (3) vs. the received signal for different  $\gamma$  for 8-PAM and the BD for  $\mathbf{p}_{85} = [0, 1, 0, 1, 0, 1, 0, 1]$ . Squares show the constellation points and crosses show the thresholds  $\beta_k$  in Theorem 6. Gray and white areas indicate  $\mathcal{Y}_0$  and  $\mathcal{Y}_1$ , resp. In (a) at  $\gamma = 5.5$  dB none of the seven thresholds is virtual. At  $\gamma \approx 4.9$  dB, shown in (b), the thresholds  $\beta_1$  and  $\beta_2$  as well as  $\beta_6$  and  $\beta_7$  become virtual, shown also in (c) for  $\rho = 4$  dB. At  $\gamma \approx 2.2$  dB, shown in (d),  $\beta_3$  and  $\beta_5$  merge with  $\beta_4$  and also become virtual.

the PBER for the BD/ABD in (8). For a given pattern  $\mathbf{p}$ , we associate the threshold  $\beta_k \in \mathcal{R}$  to the bit  $p_k$  when  $p_k \neq p_{k+1}$ . Since there is no threshold  $\beta_k$  when  $p_k = p_{k+1}$ , the number of thresholds is at most  $M - 1$ . The indices of the thresholds for the pattern  $\mathbf{p}$  form a set of indices  $\mathcal{K}$ , with  $1 \leq |\mathcal{K}| \leq M - 1$ . For example, the pattern  $\mathbf{p}_{54} = [0, 0, 1, 1, 0, 1, 1, 0]$  has  $\mathcal{K} = \{2, 4, 5, 7\}$ .

The thresholds for the ABD, which we denote by  $\tilde{\beta}$ , are independent of  $\gamma$  and placed at the midpoints between adjacent constellation points with different binary labels, which follows directly from (5). On the other hand, the thresholds for the BD depend on  $\gamma$ . We denote these thresholds by  $\beta$ , where to simplify the notation, the dependency on  $\gamma$  is omitted. Fig. 1

shows the thresholds for the BD and their dependency on the SNR. This figure shows that some thresholds can merge at low SNR and seem to disappear. To take this effect into account, we define virtual thresholds as follows. A threshold  $\beta_k$  is said to be *virtual* at  $\gamma < \gamma_0$  if it merges with another threshold  $\beta_{k'}$  at  $\gamma = \gamma_0$  (i.e.,  $\beta_k = \beta_{k'}$  when  $\gamma = \gamma_0$ ) and does not exist at  $\gamma < \gamma_0$ .

#### B. General Expression for One-dimensional Constellations

The BER expression for the ABD and an  $M$ -PAM constellation with any labeling is well known and can be found in [7, eq. (21)]. The PBER expression can easily be obtained in a similar way. The following theorem gives a general PBER

expression if none of the thresholds  $\beta_k$  is virtual.

*Theorem 1:* The PBER of the BD or the ABD using an arbitrary one-dimensional constellation with a pattern  $\mathbf{p}$  can be expressed as

$$P = \frac{1}{2} + \frac{1}{M} \sum_{i=1}^M \sum_{k \in \mathcal{K}} g_{i,k} Q\left((\beta_k - s_i)\sqrt{2\gamma}\right), \quad (9)$$

where none of  $\beta_k$  is virtual, and  $g_{i,k} \in \{\pm 1\}$  is

$$g_{i,k} \triangleq (p_{k+1} - p_k)(1 - 2p_i). \quad (10)$$

*Proof:* The proof is given in Appendix A.  $\square$

The following theorem shows that Theorem 1 also holds when some of the thresholds become virtual, provided that their values are chosen properly.

*Theorem 2:* If  $\beta_k$  is virtual for  $\gamma < \gamma_0$  because  $\beta_k = \beta_{k'}$  at  $\gamma = \gamma_0$  for some  $k$  and  $k' > k$ , the PBER for the BD is given by (9)–(10) if  $\beta_k = \beta_{k'}$ ,  $\forall \gamma < \gamma_0$ .

*Proof:* We will show that the PBER in (9) is not affected by the virtual threshold  $\beta_k$  if  $\beta_k = \beta_{k'}$ . Let  $S_i$  be the two terms in the inner sum in (9) associated to the thresholds  $\beta_k$  and  $\beta_{k'}$ , i.e.,

$$S_i \triangleq g_{i,k} Q\left((\beta_k - s_i)\sqrt{2\gamma}\right) + g_{i,k'} Q\left((\beta_{k'} - s_i)\sqrt{2\gamma}\right). \quad (11)$$

Since  $\beta_k$  and  $\beta_{k'}$  are thresholds that merge,  $p_{k+1} = p_{k'}$  and  $p_k = p_{k'+1}$  must hold. Using these relations in (10), we obtain  $g_{i,k} = -g_{i,k'}$ , which used in (11) gives  $S_i = 0$ ,  $\forall i$  if  $\beta_k = \beta_{k'}$ .  $\square$

*Remark 1:* Theorem 2 holds regardless of whether  $\beta_{k'}$  is virtual or not. If  $\beta_{k'}$  is not virtual for  $\gamma < \gamma_0$ ,  $\beta_k$  must be set to the value of  $\beta_{k'}$  for  $\gamma < \gamma_0$ . If  $\beta_{k'}$  is virtual for  $\gamma < \gamma_0$ ,  $\beta_k = \beta_{k'}$  can be set to any real value.

### C. BER for the ABD and M-PAM

For M-PAM and the ABD, (9) can be expressed as a bit-wise version of [7, eq. (21)]:

$$\tilde{P} = \frac{1}{M} \sum_{n=1}^{M-1} a_n Q\left((2n-1)d\sqrt{2\gamma}\right), \quad (12)$$

where

$$a_n = \sum_{k=n}^{M-1} (p_{k+1} - p_k)(1 - 2p_{k+1-n}) - (p_{k+2-n} - p_{k+1-n})(1 - 2p_{k+1}). \quad (13)$$

One direct consequence of (12) is that the vector  $\mathbf{a} \triangleq [a_1, \dots, a_{M-1}]$  with  $a_n$  given by (13) completely defines the performance of the ABD for M-PAM and allows us to compare the performance of different patterns. From (12), the PBER for high SNR is determined by the coefficient multiplying the Q-function with the smallest argument, i.e.,  $a_1$ . If two patterns have identical coefficients  $a_1$ , the next coefficients  $a_2$  should be compared, and so on.

Using (6) and (12), the average BER for an M-PAM with

a labeling  $\mathbb{C}$  can be expressed as

$$\tilde{P}_{\mathbb{C}} = \frac{1}{mM} \sum_{n=1}^{M-1} \alpha_n Q\left((2n-1)d\sqrt{2\gamma}\right), \quad (14)$$

where  $\boldsymbol{\alpha} \triangleq [\alpha_1, \dots, \alpha_{M-1}]$  is the sum of vectors  $\mathbf{a}$  for the  $m$  patterns used in  $\mathbb{C}$ . The equation in (14) in fact corresponds to [7, eq. (21)], where the value of  $\alpha_n$  is a scaled version of the so-called differential average distance spectrum  $\bar{\delta}(n, \lambda)$ , i.e.,  $\alpha_n = 2M\bar{\delta}(n, \lambda)$ .

### IV. BIT PATTERNS

For 4-PAM, the patterns  $\mathbf{p}_5 = [0, 1, 0, 1]$  and  $\mathbf{p}_{10} = [1, 0, 1, 0]$  have identical PBER performance because of the symmetry of the constellation. To find all the patterns with different performance, we group patterns into classes, where all the patterns in the same class have identical PBER.

In the following, we define two operations that can be applied to a pattern. A *reflection* of  $\mathbf{p}$  is defined as  $\mathbf{p}' = \text{refl}(\mathbf{p})$  with  $p'_i = p_{M+1-i}$  for  $i = 1, \dots, M$ . For example,  $\mathbf{p}_{27} = [0, 0, 0, 1, 1, 0, 1, 1] = \text{refl}([1, 1, 0, 1, 1, 0, 0, 0]) = \text{refl}(\mathbf{p}_{216})$ . A *negation* of  $\mathbf{p}$  is defined as  $\mathbf{p}' = \text{neg}(\mathbf{p})$  with  $p'_i = \bar{p}_i$  for  $i = 1, \dots, M$ . For example,  $\mathbf{p}_{39} = [0, 0, 1, 0, 0, 1, 1, 1] = \text{neg}([1, 1, 0, 1, 1, 0, 0, 0]) = \text{neg}(\mathbf{p}_{216})$ . Both these functions are self-inverse, i.e.,  $\mathbf{p} = \text{refl}(\text{refl}(\mathbf{p}))$  and  $\mathbf{p} = \text{neg}(\text{neg}(\mathbf{p}))$ , and they commute, i.e.,  $\text{refl}(\text{neg}(\mathbf{p})) = \text{neg}(\text{refl}(\mathbf{p}))$ . Note also that for any pattern  $\mathbf{p}$ , we have that  $\mathbf{p} \neq \text{neg}(\mathbf{p})$ . Using the previous definitions, we now define three special types of patterns that will be useful throughout this paper.

The pattern  $\mathbf{p}$  is said to be *reflected* (RE) if  $\text{refl}(\mathbf{p}) = \mathbf{p}$ . For example,  $\mathbf{p}_{60} = [0, 0, 1, 1, 1, 1, 0, 0]$  is an RE pattern. The pattern  $\mathbf{p}$  is said to be *anti-reflected* (ARE) if  $\text{neg}(\text{refl}(\mathbf{p})) = \mathbf{p}$ . For example,  $\mathbf{p}_{43} = [0, 0, 1, 0, 1, 0, 1, 1]$  is an ARE pattern. The pattern  $\mathbf{p}$  is called *asymmetric* (ASY) if it is neither RE nor ARE. For example,  $\mathbf{p}_{216} = [1, 1, 0, 1, 1, 0, 0, 0]$  is an ASY pattern.

From (7)–(8), we note that the PBER is not affected by reflections and/or negation of the patterns, since the PBER is averaged over both transmitted zeros and ones. Because of this, we group all patterns that are connected via reflection or negation into one class of patterns with identical PBER. Each class contains at the least two patterns ( $\mathbf{p}$  and  $\text{neg}(\mathbf{p})$ ) because  $\mathbf{p} \neq \text{neg}(\mathbf{p})$ ,  $\forall \mathbf{p}$  and at most 4 different patterns ( $\mathbf{p}$ ,  $\text{neg}(\mathbf{p})$ ,  $\text{refl}(\mathbf{p})$ , and  $\text{neg}(\text{refl}(\mathbf{p}))$ ) if they are all different and is represented by a unique class index  $q \in 1, \dots, Q$ , where  $Q$  is the number of classes. A labeling can now be represented not only by a binary matrix  $\mathbb{C}$  or by the set of pattern indices  $\mathcal{W}$ , but also by an ordered vector of class indices  $\mathbf{q} = [q_1, q_2, \dots, q_m] \in \{1, \dots, Q\}^m$ , where  $q_1 \leq q_2 \leq \dots \leq q_m$ . The reason for introducing this vector  $\mathbf{q}$  is that it allows us to easily identify two binary labelings that give the same BER. In other words, if two different labelings  $\mathbb{C}$  and  $\mathbb{C}'$  have vectors  $\mathbf{q}$  and  $\mathbf{q}'$ , resp., they will give the same average BER if  $\mathbf{q} = \mathbf{q}'$ . To clarify these definitions, consider the following example (see also Example 2).

*Example 1 (Patterns for 4-PAM):* For 4-PAM there are six patterns which are grouped into  $Q = 3$  classes as shown in Table I. Two of them are RE and four are ARE, as indicated

TABLE I

CLASSES OF PATTERNS FOR 4-PAM WITH THEIR CORRESPONDING CLASS INDICES  $q$ , REPRESENTATIVES  $\mathbf{p}$ , TYPES, DECIMAL REPRESENTATIONS OF THE PATTERNS  $w$ , VECTORS  $\mathbf{a}$  DEFINING THE PBER FOR THE ABD, AND THRESHOLDS FOR THE REPRESENTATIVES

$q$	$\mathbf{p}$	Type	$w$	$\mathbf{a}$	Thresholds
1	[0, 0, 1, 1]	ARE	<b>3</b> 12	[2, 2, 0]	$\beta_2 = 0$
2	[0, 1, 1, 0]	RE	<b>6</b> 9	[4, 2, -2]	$\beta_3 = -\beta_1$
3	[0, 1, 0, 1]	ARE	<b>5</b> 10	[6, -4, 2]	$\beta_3 = -\beta_1, \beta_2 = 0$

in the third column of Table I. The first column is the index of the class and the fourth column contains the decimal representations of the indices of the patterns that belong to that class. As an example, one of the patterns of a class is shown in the second column. These patterns are called *representative* and are used to analyze the performance of the patterns in the class. The indices of the representatives are shown with boldface in column 4. The fifth column contains vectors  $\mathbf{a}$  defining the PBER for the ABD in (12). The patterns are ordered from best to worst PBER for high SNR, as predicted by the vectors  $\mathbf{a}$ . All  $4! = 24$  valid labelings for 4-PAM give three vectors  $\mathbf{q}$ , i.e., three labelings that give different BER:  $\mathbf{q} = [1, 2]$  (BRGC),  $\mathbf{q} = [1, 3]$  (NBC), and  $\mathbf{q} = [2, 3]$  (AGC). These labelings are listed in the first part of Table II together with their class indices  $q$ , pattern indices  $\mathcal{W}$ , and vectors  $\alpha$  defining the BER for the ABD. The labelings are also ordered from best to worst BER for high SNR as predicted by the vectors  $\alpha$ .

The next theorem gives closed form expressions for the number of classes for length- $M$  patterns.

*Theorem 3:* For  $M$ -PAM, all the length- $M$  bit patterns can be grouped into  $Q$  classes, where

$$Q = \frac{1}{4} \left( \binom{M}{M/2} + \binom{M/2}{M/4} + 2^{M/2} \right), \quad (15)$$

among which  $Q_{\text{RE}}$  classes have only RE patterns,  $Q_{\text{ARE}}$  only ARE patterns, and  $Q_{\text{ASY}}$  only asymmetric patterns, where

$$Q_{\text{RE}} = \frac{1}{2} \binom{M/2}{M/4}, \quad (16)$$

$$Q_{\text{ARE}} = 2^{M/2-1}, \quad (17)$$

$$Q_{\text{ASY}} = \frac{1}{4} \left( \binom{M}{M/2} - \binom{M/2}{M/4} - 2^{M/2} \right). \quad (18)$$

*Proof:* Any pattern  $\mathbf{p}$  must contain  $M/2$  zeros and  $M/2$  ones, hence, the total number of patterns is equal to  $\binom{M}{M/2}$ . For a pattern to be RE,  $p_i = p_{M-i+1}$ , i.e., the positions of the  $M/4$  ones in  $[p_1, \dots, p_{M/2}]$  fully describe the pattern, and thus, the number of RE patterns is  $\binom{M/2}{M/4}$ . There are two members in every class of RE patterns,  $\mathbf{p} = \text{refl}(\mathbf{p})$  and  $\text{neg}(\text{refl}(\mathbf{p})) = \text{neg}(\mathbf{p})$ , which gives (16).

For a pattern to be ARE,  $p_i = \bar{p}_{M-i+1}$ , i.e., the positions of the ones in  $[p_1, \dots, p_{M/2}]$  fully describe the pattern where the number of ones in  $[p_1, \dots, p_{M/2}]$  is between 0 and  $M/2$ . From that, it follows that there are  $2^{M/2}$  ARE patterns. There are two members in every class of ARE patterns ( $\mathbf{p} = \text{neg}(\text{refl}(\mathbf{p}))$  and  $\text{refl}(\mathbf{p}) = \text{neg}(\mathbf{p})$ ), which gives (17).

All the remaining classes include only asymmetric patterns. The number of asymmetric patterns can be obtained by sub-

TABLE II

SOME COMMON LABELINGS FOR 4-PAM AND 8-PAM WITH THEIR CORRESPONDING CLASS INDICES  $q$ , PATTERNS INDICES  $\mathcal{W}$ , AND VECTORS  $\alpha$  DEFINING THE BER FOR THE ABD

$M$	Labeling	$\mathbf{q}$	$\mathcal{W}$	$\alpha$
4	BRGC	[1, 2]	{3, 6}	[6, 4, -2]
4	NBC	[1, 3]	{3, 5}	[8, -2, 2]
4	AGC	[2, 3]	{5, 6}	[10, -2, 0]
8	BRGC	[1, 2, 6]	{15, 60, 102}	[14, 12, -2, 0, 2, 0, -2]
8	FBC	[1, 2, 10]	{15, 60, 90}	[18, 0, 4, 10, -8, 2, -2]
8	NBC	[1, 5, 11]	{15, 51, 85}	[22, -4, 8, -10, 8, -2, 2]
8	BSGC	[2, 6, 9]	{105, 60, 102}	[22, 10, -8, 4, -2, -2, 0]
8	AGC	[9, 10, 11]	{90, 105, 85}	[36, -18, 6, 4, -4, -2, 2]

tracting  $2Q_{\text{RE}}$  and  $2Q_{\text{ARE}}$  from the total number of patterns, i.e.,  $\binom{M}{M/2} - 2Q_{\text{RE}} - 2Q_{\text{ARE}}$ . There are four patterns in each class, as  $\mathbf{p} \neq \text{refl}(\mathbf{p})$  and  $\mathbf{p} \neq \text{refl}(\text{neg}(\mathbf{p}))$  (or equivalently,  $\text{refl}(\mathbf{p}) \neq \text{neg}(\mathbf{p})$ ). Using this, (18) is obtained.

Finally, the total number of classes in (15) is obtained as  $Q_{\text{RE}} + Q_{\text{ARE}} + Q_{\text{ASY}}$ .  $\square$

Theorem 3 gives the exact number of classes  $Q$  for  $M$ -PAM constellations. A loose bound on this number of classes was previously presented in [25, eq. (3.14)]. For 4-PAM Theorem 3 gives  $Q = 3$ ,  $Q_{\text{RE}} = 1$ ,  $Q_{\text{ARE}} = 2$ , and  $Q_{\text{ASY}} = 0$ , which is in agreement with Example 1 and Table I. The following two examples show the number of classes for 8-PAM and 16-PAM.

*Example 2 (Patterns for 8-PAM):* For 8-PAM ( $M = 8$ ), Theorem 3 states that there are  $Q = 23$  classes of patterns,  $Q_{\text{RE}} = 3$  classes of RE patterns and  $Q_{\text{ARE}} = 8$  classes of ARE patterns, each containing two patterns. These classes are shown in the first 11 rows of Table III and ordered from the best to worst PBER for high SNR, as predicted by the vectors  $\mathbf{a}$ . The  $Q_{\text{ASY}} = 12$  classes of asymmetric patterns, each with four members, are listed in the last 12 rows of Table III and also ordered in a similar way. By enumerating all  $8! = 40320$  valid labelings for 8-PAM, we found 460 vectors  $\mathbf{q}$ , i.e., for 8-PAM there exist only 460 labelings that give different BER. The classes associated to five of the most well-known labelings are:  $\mathbf{q} = [1, 2, 6]$  (BRGC),  $\mathbf{q} = [1, 2, 10]$  (FBC),  $\mathbf{q} = [1, 5, 11]$  (NBC),  $\mathbf{q} = [2, 6, 9]$  (BSGC), and  $\mathbf{q} = [9, 10, 11]$  (AGC). These labelings are listed from best to worst BER in the second part of Table II.

*Example 3 (Patterns for 16-PAM):* For 16-PAM ( $M = 16$ ) Theorem 3 states that there are  $Q = 3299$  classes,  $Q_{\text{RE}} = 35$  of the classes contain RE patterns,  $Q_{\text{ARE}} = 128$  classes contain ARE patterns, and  $Q_{\text{ASY}} = 3136$  classes contain ASY patterns. The patterns for 16-PAM are not listed in this paper.

## V. THRESHOLDS FOR THE BD

In this section, we show that the thresholds for the BD can be found by solving an  $(M-1)$ th power polynomial equation and give a closed-form solutions for 4-PAM and 8-PAM with RE and ARE patterns.

### A. Threshold Computation

The problem of finding the thresholds  $\beta_k$  for the BD is equivalent to finding the solutions of  $l(y) = 0$ . In the

TABLE III  
CLASSES OF PATTERNS FOR 8-PAM WITH THEIR CORRESPONDING CLASS INDICES  $q$ , REPRESENTATIVES  $\mathbf{p}$ , TYPES, DECIMAL REPRESENTATIONS OF THE PATTERNS  $w$ , VECTORS  $\mathbf{a}$  DEFINING THE PBER FOR THE ABD, AND THRESHOLDS FOR THE REPRESENTATIVES

$q$	$\mathbf{p}$	Type	$w$	$\mathbf{a}$	Thresholds
1	[0, 0, 0, 0, 1, 1, 1, 1]	ARE	<b>15</b> 240	[ 2, 2, 2, 2, 0, 0, 0, 0]	$\beta_4 = 0$
2	[0, 0, 1, 1, 1, 1, 0, 0]	RE	<b>60</b> 195	[ 4, 4, 2, 2, -2, -2, 0, 0]	$\beta_6 = -\beta_2 = f(t_2)$
3	[1, 1, 1, 0, 1, 0, 0, 0]	ARE	23 <b>232</b>	[ 6, -2, 2, 0, 2, 0, 0, 0]	$\beta_5 = -\beta_3 = f(t_1), \beta_4 = 0$
4	[0, 1, 1, 1, 0, 0, 0, 1]	ARE	<b>113</b> 142	[ 6, 4, 4, -4, -2, -2, 2, 2]	$\beta_7 = -\beta_1 = f(t_2), \beta_4 = 0$
5	[0, 0, 1, 1, 0, 0, 1, 1]	ARE	<b>51</b> 204	[ 6, 6, -4, -4, 2, 2, 0, 0]	$\beta_6 = -\beta_2 = f(t_2), \beta_4 = 0$
6	[0, 1, 1, 0, 0, 1, 1, 0]	RE	<b>102</b> 153	[ 8, 6, -6, -4, 4, 2, -2, 0]	$\beta_7 = -\beta_1 = f(t_2), \beta_5 = -\beta_3 = f(t_3)$
7	[0, 0, 1, 0, 1, 0, 1, 1]	ARE	<b>43</b> 212	[10, -6, 4, -2, 0, 2, 0, 0]	$\beta_6 = -\beta_2 = f(t_2),$ $\beta_5 = -\beta_3 = f(t_3), \beta_4 = 0$
8	[0, 1, 0, 0, 1, 1, 0, 1]	ARE	<b>77</b> 178	[10, 0, -6, 2, 4, -4, 2, 2]	$\beta_7 = -\beta_1 = f(t_2),$ $\beta_6 = -\beta_2 = f(t_3), \beta_4 = 0$
9	[0, 1, 1, 0, 1, 0, 0, 1]	ARE	<b>105</b> 150	[10, 0, -4, 6, -4, -2, 2, 2]	$\beta_7 = -\beta_1 = f(t_2),$ $\beta_5 = -\beta_3 = f(t_3), \beta_4 = 0$
10	[1, 0, 1, 0, 0, 1, 0, 1]	RE	90 <b>165</b>	[12, -6, 0, 6, -6, 4, -2, 2]	$\beta_7 = -\beta_1 = f(t_1),$ $\beta_6 = -\beta_2 = f(t_3), \beta_5 = -\beta_3 = f(t_2)$
11	[0, 1, 0, 1, 0, 1, 0, 1]	ARE	<b>85</b> 170	[14, -12, 10, -8, 6, -4, 2, 2]	$\beta_7 = -\beta_1 = f(t_2), \beta_6 = -\beta_2 = f(t_3),$ $\beta_5 = -\beta_3 = f(t_1), \beta_4 = 0$
12	[0, 0, 0, 1, 1, 1, 1, 0]	ASY	<b>30</b> 120 135 225	[ 4, 3, 3, 2, -2, -1, -1]	$\beta_3, \beta_7$
13	[0, 0, 0, 1, 1, 1, 0, 1]	ASY	<b>29</b> 71 184 226	[ 6, 1, 2, -3, 1, 0, 1]	$\beta_3, \beta_6, \beta_7$
14	[0, 0, 0, 1, 1, 0, 1, 1]	ASY	<b>27</b> 39 216 228	[ 6, 2, -3, 1, 1, 1, 0]	$\beta_3, \beta_5, \beta_6$
15	[0, 0, 1, 1, 1, 0, 0, 1]	ASY	<b>57</b> 99 156 198	[ 6, 5, 0, -3, -3, 2, 1]	$\beta_2, \beta_5, \beta_7$
16	[0, 0, 1, 0, 1, 1, 1, 0]	ASY	<b>46</b> 116 139 209	[ 8, -1, 2, -1, 3, -2, -1]	$\beta_2, \beta_3, \beta_4, \beta_7$
17	[0, 0, 1, 1, 1, 0, 1, 0]	ASY	<b>58</b> 92 163 197	[ 8, -1, 3, -2, 2, -1, -1]	$\beta_2, \beta_5, \beta_6, \beta_7$
18	[0, 1, 0, 0, 1, 1, 1, 0]	ASY	<b>78</b> 114 141 177	[ 8, 2, -1, -1, -1, 3, -2]	$\beta_1, \beta_2, \beta_4, \beta_7$
19	[0, 0, 1, 1, 0, 1, 1, 0]	ASY	<b>54</b> 108 147 201	[ 8, 3, -6, 3, 3, -2, -1]	$\beta_2, \beta_4, \beta_5, \beta_7$
20	[0, 0, 1, 0, 1, 1, 1, 0]	ASY	<b>45</b> 75 180 210	[10, -3, -3, 6, -4, 1, 1]	$\beta_2, \beta_3, \beta_4, \beta_6, \beta_7$
21	[0, 0, 1, 1, 0, 1, 0, 1]	ASY	<b>53</b> 83 172 202	[10, -3, 1, 0, -2, 1, 1]	$\beta_2, \beta_4, \beta_5, \beta_6, \beta_7$
22	[0, 1, 0, 1, 1, 0, 0, 1]	ASY	<b>89</b> 101 154 166	[10, 0, -3, 1, 1, -3, 2]	$\beta_1, \beta_2, \beta_3, \beta_5, \beta_7$
23	[0, 1, 0, 1, 0, 1, 1, 0]	ASY	<b>86</b> 106 149 169	[12, -6, 3, -1, -1, 3, -2]	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_7$

following theorem we show how this can be done for  $M$ -PAM constellations.

**Theorem 4:** The thresholds for the BD and  $M$ -PAM constellations with a pattern  $\mathbf{p}$  are

$$\beta_k = \frac{1}{4\gamma d} \log z_n, \quad (19)$$

where  $z_n$  are the real and positive solutions of

$$\sum_{i=1}^M \check{p}_i A^{\frac{(M/2-i)(M/2+1-i)}{2}} z^{i-1} = 0, \quad (20)$$

and

$$A = e^{-8\gamma d^2}. \quad (21)$$

*Proof:* Using (3),  $l(y) = 0$  is equivalent to  $e^{l(y)} = 1$ , which can be restated as

$$h(y) \triangleq \sum_{i=1}^M \check{p}_i e^{-\gamma(y+d(M-2i+1))^2}, \quad (22)$$

where the definition of the  $M$ -PAM symbols was used. Factorizing (22) gives

$$h(y) = e^{-\gamma y^2 - 2\gamma d y(M-1) - \gamma d^2} \cdot \sum_{i=1}^M \check{p}_i e^{4\gamma d y(i-1)} e^{-8\gamma d^2 \frac{(M/2-i)(M/2+1-i)}{2}}. \quad (23)$$

Using (21) in (23) together with substitution  $z = e^{4\gamma d y}$ , (20) is obtained by setting  $h(y) = 0$  and removing the nonzero factor preceding the summation in (23). The expression in (20) is an  $(M-1)$ -power polynomial<sup>1</sup>, and thus, it always has  $M-1$  solutions. Because of the substitution  $z = e^{4\gamma d y}$ , only the positive (and real) roots need to be considered.  $\square$

Theorem 4 gives a general expression for finding the thresholds for  $M$ -PAM with any pattern  $\mathbf{p}$ . After finding the roots of (20), the thresholds  $\beta_k$  may easily be obtained from (19). The main problem is that finding the roots of (20) does not admit simple closed-form solutions. However, it can always be solved numerically. Fig. 2 illustrates the thresholds for 16-PAM with the pattern  $\mathbf{p}_{45745} =$

<sup>1</sup>Interestingly, for  $i = 1, \dots, M-1$  with  $i \neq M/2$  and  $i \neq M/2+1$ , the powers of  $A$  in (20) are the so-called “triangular numbers”.

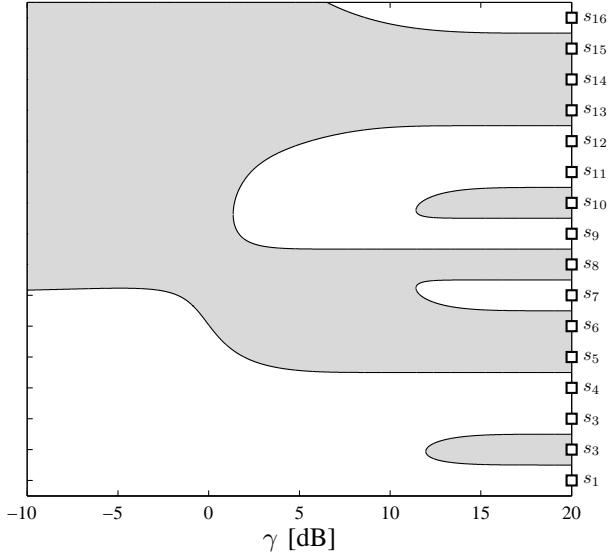


Fig. 2. Thresholds  $\beta_k$  in (19) for 16-PAM with the ASY pattern  $\mathbf{p}_{45745} = [1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1]$  obtained by solving (20) numerically. The constellation points are shown with squares. Gray and white areas indicate  $\mathcal{Y}_0$  and  $\mathcal{Y}_1$ , resp.

$[1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1]$  obtained by solving (20) numerically. In the following two sections we show how this problem can be solved analytically for 4-PAM and 8-PAM with RE or ARE patterns.

### B. Thresholds for 4-PAM

The following theorem shows how the thresholds are found for 4-PAM with any pattern.

**Theorem 5:** The thresholds  $\beta_k$  for any pattern for 4-PAM, as listed in the last column in Table I, can be expressed as

$$\beta_1 = -\beta_3, \quad (24)$$

$$\beta_2 = 0, \quad (25)$$

$$\beta_3 = \frac{1}{4\gamma d} \log \left| \frac{1 + \check{p}_1 \check{p}_4 A + \sqrt{(1 + \check{p}_1 \check{p}_4 A)^2 - 4A^2}}{2A} \right|, \quad (26)$$

where  $A$  is given by (21).

*Proof:* The proof is given in Appendix B.  $\square$

Theorem 5 gives closed-form expressions for any pattern for 4-PAM, and thus, it allows us to compute the BER for all the labelings in the first part of Table II. The results in Theorem 5 can be shown to coincide to those in [4, eq. (10)] when the BRGC is considered.

### C. Thresholds for 8-PAM

The following theorem shows how the thresholds may be found for 8-PAM with RE or ARE patterns. These patterns are of great value because all the most commonly studied labelings (e.g., BRGC, NBC, FBC, BSGC, AGC) can be composed from them (cf. Table II).

**Theorem 6:** The thresholds  $\beta_k$  for the patterns in the classes

$q = 1, 2, \dots, 11$  for 8-PAM can be expressed as

$$\beta_k = -\beta_{8-k} = \begin{cases} f(t_n) & \text{if } k = 5, 6, 7, \\ 0 & \text{if } k = 4, \end{cases} \quad (27)$$

$$f(t) \triangleq \frac{1}{4\gamma d} \log \left| \frac{|t| + \sqrt{|t|^2 - 4}}{2} \right|, \quad (28)$$

with

$$t_1 \triangleq \frac{1}{6\check{p}_1 A^3} \left( T + \frac{2\sqrt[3]{2}C}{B} + \sqrt[3]{4B} \right), \quad (29)$$

$$t_2 \triangleq \frac{1}{6\check{p}_1 A^3} \left( T - \frac{\sqrt[3]{2}(1 + \sqrt{3}j)C}{B} - \frac{1 - \sqrt{3}j}{\sqrt[3]{2}} B \right), \quad (30)$$

$$t_3 \triangleq \frac{1}{6\check{p}_1 A^3} \left( T - \frac{\sqrt[3]{2}(1 - \sqrt{3}j)C}{B} - \frac{1 + \sqrt{3}j}{\sqrt[3]{2}} B \right), \quad (31)$$

$A$  is given by (21),

$$T \triangleq 2(\check{p}_8 A^3 - \check{p}_2),$$

$$B \triangleq \sqrt[3]{\sqrt{D^2 - 4C^3} - \check{p}_1 \check{p}_8 D},$$

$$C \triangleq 7A^6 + \check{p}_2 \check{p}_8 A^3 - 3\check{p}_1 \check{p}_3 A + 1,$$

$$D \triangleq 7\check{p}_1 A^9 - 12\check{p}_1 \check{p}_2 \check{p}_8 A^6 - 18\check{p}_3 A^4$$

$$+ 3\check{p}_1(1 + 9\check{p}_4 \check{p}_8) A^3 - 9\check{p}_2 \check{p}_3 \check{p}_8 A + 2\check{p}_1 \check{p}_2 \check{p}_8,$$

and the relationship between  $k$  and  $n$  in (27) for the different classes  $q$  is listed in the last column of Table III. As the relationship between  $n$  and  $k$  depends on the particular pattern, the representative of the class (the second column of Table III) should be used in the presented equations.

*Proof:* The proof is given in Appendix C.  $\square$

Theorem 6 shows how to analytically obtain the thresholds for the BD with 8-PAM and any RE or ARE pattern, for instance, thresholds shown in Fig. 1. Using these results, the PBER can be calculated using (9), which gives PBER expressions for 11 out of 23 classes, or equivalently, for 56 different labelings, including the 5 shown in the second part of Table II.

**Remark 2:** In the high SNR regime, i.e.,  $\gamma \rightarrow \infty$ , all the thresholds in Theorems 5 and 6 tend to midpoints, i.e., the same constant thresholds used in the ABD for all SNR. This fact can easily be proven analytically for 4-PAM by evaluating  $\lim_{\gamma \rightarrow \infty} \beta_3$  and applying l'Hôpital's rule. For 8-PAM a similar proof exists, however, in this case it is not straightforward due to the complexity of the threshold expressions. These results can be intuitively understood from the fact that the max-log approximation in (5) becomes more precise when the SNR increases, and hence, the thresholds for the BD and ABD are expected to coincide when  $\gamma \rightarrow \infty$ .

## VI. NUMERICAL RESULTS

In Fig. 3 we show the thresholds given by Theorem 6 for the pattern  $\mathbf{p}_{165}$  ( $q = 10$ ) for 8-PAM. The figure is symmetric with respect to zero due to the symmetry of the pattern. At  $\gamma \approx 5.3$  dB the pairs of thresholds  $\beta_2$  and  $\beta_3$ , and  $\beta_5$  and  $\beta_6$  merge and become virtual for all  $\gamma < 5.3$  dB. All the virtual thresholds shown with dashed lines satisfy the conditions in

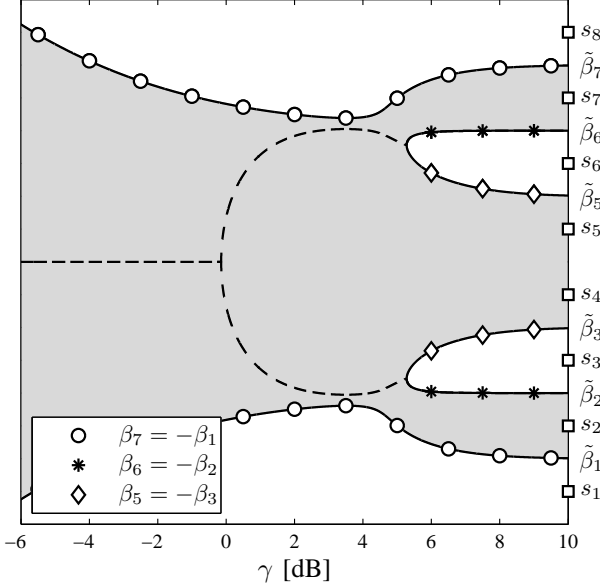


Fig. 3. The thresholds for  $\mathbf{p}_{165} = [1, 0, 1, 0, 0, 1, 0, 1]$  ( $q = 10$ ) for 8-PAM in Theorem 6 vs. SNR. Virtual thresholds are shown with dashed lines. The thresholds for the ABD  $\tilde{\beta}_k$  and the constellation points (squares) are also shown. Gray and white areas indicate  $\mathcal{Y}_0$  and  $\mathcal{Y}_1$ , resp.

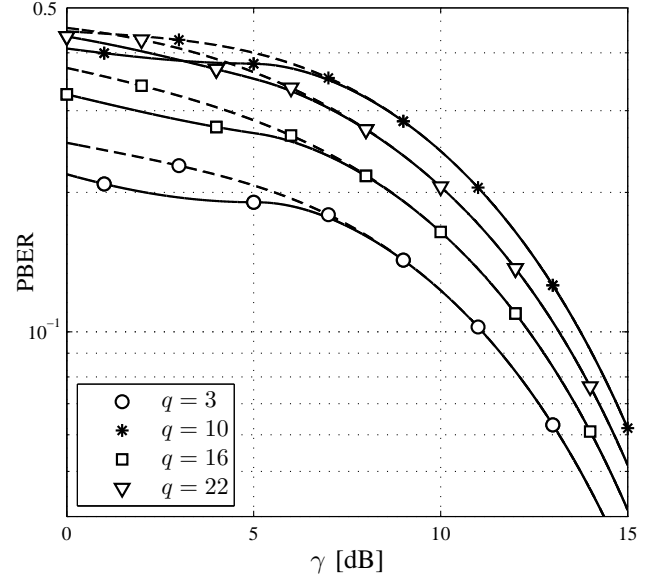


Fig. 4. The PBBER for 8-PAM with ARE ( $q = 3$ ) and RE ( $q = 10$ ) patterns and ASY patterns ( $q = 16, 22$ ). Solid lines correspond to the BD and dashed lines to the ABD. The threshold for the BD were obtained using Theorem 6 for  $q = 3, 10$  (the thresholds for  $q = 10$  are shown in Fig. 3) and solving (20) numerically for  $q = 16, 22$ .

Theorem 2. As expected (see Remark 2), when  $\gamma \rightarrow \infty$ , the BD thresholds coincide with the ABD thresholds.

The PBBER for 8-PAM with some selected patterns from Table III using (9) is presented in Fig. 4. The thresholds are calculated analytically for  $q = 3, 10$  and numerically for  $q = 16, 22$ . For very low SNR the gap between the BD and the ABD can reach up to several dB, however, this gap decreases when the SNR increases. The same conclusion can be drawn for all other patterns except for  $q = 1$ , as in this case only one threshold exists  $\beta_{M/2} = 0$  for all SNR. Hence, for  $q = 1$  the BD and the ABD have the same performance for  $M$ -PAM. To conclude, we present in Fig. 5 the BER for 8-PAM with the labelings in Table II. From the presented results we conclude that the BD outperforms the ABD, however, for any BER of practical interest (below 0.1), the difference between the BD and the ABD is negligible.

## VII. CONCLUSIONS

We proposed a general approach for estimating the performance of the optimal bit-wise demodulator and presented closed-form expressions for the BER for 4-PAM and 8-PAM with different labelings. We conclude that a suboptimal symbol-wise demodulator shows no loss compared to the optimal demodulator for all the SNR of interest, which justifies its use in practical systems.

The derived BER expressions for the optimal demodulator can be used to calculate the mutual information (MI) of BICM when the demodulator makes hard decisions on the bits. The optimal bit-wise demodulator does not necessarily maximize the MI. Finding the hard-decision demodulator that maximizes the MI is left for future work.

The proposed technique for finding the zero crossings of the L-values for 8-PAM works only for reflected or anti-reflected

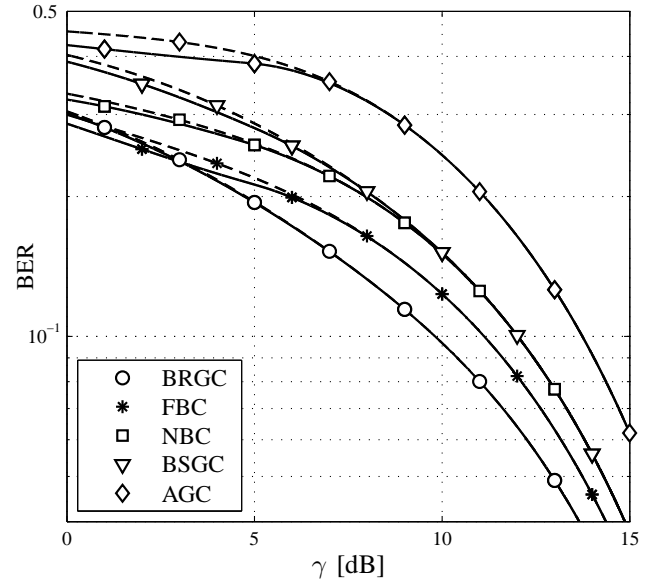


Fig. 5. The BER for 8-PAM with the 5 labelings in Table II. Solid lines correspond to the BD and dashed lines to the ABD.

patterns, which includes 11 out of 23 classes of patterns. Extending these results to the remaining classes of patterns for 8-PAM is left for further investigation as well as generalizing the results to arbitrary  $M$ .



## APPENDIX A PROOF OF THEOREM 1

Let  $v_{i,k}$  be the conditional probabilities

$$v_{i,1} \triangleq \Pr\{Y \leq \beta_1 | X = s_i\} \\ = 1 - Q\left((\beta_1 - s_i)\sqrt{2\gamma}\right), \quad (32)$$

$$v_{i,k} \triangleq \Pr\{\beta_{k-1} < Y \leq \beta_k | X = s_i\} \\ = Q\left((\beta_{k-1} - s_i)\sqrt{2\gamma}\right) - Q\left((\beta_k - s_i)\sqrt{2\gamma}\right), \quad (33)$$

$$v_{i,M} \triangleq \Pr\{\beta_{M-1} < Y | X = s_i\} \\ = Q\left((\beta_{M-1} - s_i)\sqrt{2\gamma}\right), \quad (34)$$

where  $\beta_k$  for  $k \in \mathcal{K}$  are the thresholds and none of them is virtual. The PBER in (8) can now be rewritten as

$$P = \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^M e_{i,k} v_{i,k}, \quad (35)$$

where  $e_{i,k} \triangleq p_i \oplus p_k \in \{0, 1\}$ .

Using (32)–(34) the PBER in (35) can be expressed as

$$P = \frac{1}{M} \left[ \sum_{i=1}^M e_{i,1} + \sum_{i=1}^M \sum_{k=2}^M e_{i,k} Q\left((\beta_{k-1} - s_i)\sqrt{2\gamma}\right) \right. \\ \left. - \sum_{i=1}^M \sum_{k=1}^{M-1} e_{i,k} Q\left((\beta_k - s_i)\sqrt{2\gamma}\right) \right] \\ = \frac{1}{2} + \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^{M-1} (e_{i,k+1} - e_{i,k}) Q\left((\beta_k - s_i)\sqrt{2\gamma}\right), \quad (36)$$

where  $\sum_{i=1}^M e_{i,1} = \sum_{i=1}^M p_i \oplus p_1 = M/2$  was used. To obtain the expression in (9), we express  $e_{i,k+1} - e_{i,k}$  in (36) as

$$e_{i,k+1} - e_{i,k} = p_{k+1} \oplus p_i - p_k \oplus p_i \quad (37)$$

$$= (p_{k+1} - p_k)(1 - 2p_i), \quad (38)$$

where the identity  $p_i \oplus p_j = p_i \bar{p}_j + \bar{p}_i p_j$  was used together with  $\bar{p}_i = 1 - p_i$ .

## APPENDIX B PROOF OF THEOREM 5

Define the function  $h(z)$  as

$$h(z) = \check{p}_4 A z^3 + \check{p}_3 z^2 + \check{p}_2 z + \check{p}_1 A. \quad (39)$$

According to Theorem 4, for 4-PAM with a pattern  $\mathbf{p} = [p_1, p_2, p_3, p_4]$ , equation  $h(z) = 0$  needs to be solved in order to find the thresholds. The patterns for 4-PAM are either RE ( $\check{p}_i = \check{p}_{M+1-i}$ ,  $\forall i$ ) or ARE ( $\check{p}_i = -\check{p}_{M+1-i}$ ,  $\forall i$ ). Therefore

$$h(z) = \check{p}_1 A z^3 + \check{p}_2 z^2 \pm \check{p}_2 z \pm \check{p}_1 A, \quad (40)$$

where the upper and the lower signs correspond to RE and ARE patterns, resp. Using  $\check{p}_i^2 = 1$  and the fact that  $\check{p}_1 \check{p}_4 = \pm 1$  for RE and ARE patterns, resp.,  $h(z)$  can be factorized as

$$h(z) = \check{p}_1 (z + \check{p}_1 \check{p}_4) (A z^2 + (\check{p}_1 \check{p}_2 - A \check{p}_1 \check{p}_4) z + A). \quad (41)$$

Solving  $h(z) = 0$  gives the three roots  $z_1 = -\check{p}_1 \check{p}_4$  and

$$z_{2,3} = \frac{\check{p}_1 \check{p}_4 A - \check{p}_1 \check{p}_2 \pm \sqrt{(\check{p}_1 \check{p}_4 A - \check{p}_1 \check{p}_2)^2 - 4A^2}}{2A}. \quad (42)$$

For  $q = 1$  (where  $\check{p}_1 \check{p}_4 = -1$  and  $\check{p}_1 \check{p}_2 = 1$ ) the root  $z_1 = 1$ , that used in (19) gives the threshold  $\beta_2 = 0$ . The other two roots in (42) are complex for all SNR and do not result in thresholds.

In a similar way, for  $q = 3$  (where  $\check{p}_1 \check{p}_2 = -1$  and  $\check{p}_1 \check{p}_4 = -1$ )  $\beta_2 = 0$ . When  $A \leq 1/3$ , or equivalently, when  $\gamma \geq 5 \log 3/8 \approx -1.63$  dB, the roots in (42) are real and positive resulting in thresholds  $\beta_1$  and  $\beta_3$  by using (19). When  $A > 1/3$  (low SNR), the roots in (42) are complex and can no longer be used in (19) for calculating the thresholds. To overcome this,  $|z_2|$  and  $|z_3|$  are used in the calculation of the thresholds, which together with  $\check{p}_1 \check{p}_2 = -1$  gives (26). The use of  $|\cdot|$  does not affect the result when the roots are real. When the roots are complex, their absolute values are equal to one (can be seen from (42)), which gives two virtual thresholds  $\beta_1 = \beta_3 = 0$  merging with the zero-threshold  $\beta_2$  at around  $-1.63$  dB. Theorem 2 allows the use of these thresholds in the calculation of the PBER.

Finally, for  $q = 2$  (where  $\check{p}_1 \check{p}_2 = -1$  and  $\check{p}_1 \check{p}_4 = 1$ )  $z_1 = -1$ , which results in no threshold. The two roots in (42) are positive for all SNR, resulting in the thresholds  $\beta_1$  and  $\beta_3$  by using (19). As the roots are positive the use of  $|\cdot|$  does not affect the result, which gives (26). This completes the proof.

## APPENDIX C PROOF OF THEOREM 6

Define the function  $h(z)$  as

$$h(z) = \check{p}_8 A^6 z^7 + \check{p}_7 A^3 z^6 + \check{p}_6 A z^5 + \check{p}_5 z^4 + \check{p}_4 z^3 \\ + \check{p}_3 A z^2 + \check{p}_2 A^3 z + \check{p}_1 A^6. \quad (43)$$

According to Theorem 4 for 8-PAM with pattern  $\mathbf{p} = [p_1, p_2, \dots, p_8]$  equation  $h(z) = 0$  needs to be solved in order to find the thresholds. For RE and ARE patterns,  $\check{p}_i = \pm \check{p}_{M+1-i}$ ,  $\forall i$ , where the upper and the lower signs correspond to RE and ARE patterns, resp. Using this property,  $h(z)$  for RE and ARE patterns is

$$h(z) = \check{p}_1 A^6 z^7 + \check{p}_2 A^3 z^6 + \check{p}_3 A z^5 + \check{p}_4 z^4 \\ \pm \check{p}_4 z^3 \pm \check{p}_3 A z^2 \pm \check{p}_2 A^3 z \pm \check{p}_1 A^6. \quad (44)$$

Factorizing (44) gives

$$h(z) = (z \pm 1) (\check{p}_1 A^6 z^6 + [\check{p}_2 A^3 \mp \check{p}_1 A^6] z^5 \\ + [\check{p}_1 A^6 \mp \check{p}_2 A^3 + \check{p}_3 A] z^4 + [\mp \check{p}_1 A^6 + \check{p}_2 A^3 \mp \check{p}_3 A + \check{p}_4] z^3 \\ + [\check{p}_1 A^6 \mp \check{p}_2 A^3 + \check{p}_3 A] z^2 + [\check{p}_2 A^3 \mp \check{p}_1 A^6] z + \check{p}_1 A^6). \quad (45)$$

Rearranging the terms in (45)  $h(z)$  can be written as

$$h(z) = z^3 (z \pm 1) (\check{p}_1 A^6 (z^3 + z^{-3}) + [\check{p}_2 A^3 \mp \check{p}_1 A^6] (z^2 + z^{-2}) \\ + [\check{p}_1 A^6 \mp \check{p}_2 A^3 + \check{p}_3 A] (z^1 + z^{-1}) \\ + [\mp \check{p}_1 A^6 + \check{p}_2 A^3 \mp \check{p}_3 A + \check{p}_4]). \quad (46)$$

Using the substitution

$$t(z) = z + z^{-1}, \quad (47)$$

(46) can be modified to

$$\begin{aligned} h(z) = & z^3(z \pm 1) (\check{p}_1 A^6 t(z)^3 + [\check{p}_2 A^3 \mp \check{p}_1 A^6] t(z)^2 \\ & + [-2\check{p}_1 A^6 \mp \check{p}_2 A^3 + \check{p}_3 A] t(z) \\ & + [\pm \check{p}_1 A^6 - \check{p}_2 A^3 \mp \check{p}_3 A + \check{p}_4]) . \end{aligned} \quad (48)$$

Finding positive roots of  $h(z) = 0$  can now be done analytically. For ARE patterns the second factor in (48) gives a root equal to one resulting in  $\beta_4 = 0$ . The roots of the last factor in (48) need to be found. As a first step we solve it with respect to  $t(z)$ , where the roots  $t_n$  are shown in (29)–(31), where  $\pm 1$  was replaced by  $\pm \check{p}_1 \check{p}_8$  to distinguish between RE and ARE patterns. As  $z$  should be real and positive, only real and positive  $t_n$  need to be considered. Two out of three roots  $t_n$  may combine into a complex conjugated couple, but the third root is always real. Every positive root  $t_n$  gives two roots for  $z$  in (45), which can be found from (47) as

$$z_{2n-1,2n} = \frac{t_n \pm \sqrt{t_n^2 - 4}}{2}, \quad (49)$$

where  $z_{2n-1} = 1/z_{2n}$ . Due to (19) and the symmetry of the patterns, these two roots give the two thresholds  $\beta_k = -\beta_{8-k}$ , which justifies the first equality in (27). When  $t_n$  is real and  $t_n \geq 2$ , the roots in (49) are positive and give thresholds  $\beta_k = -\beta_{8-k}$ . Because of  $t_n$  is real and the roots  $z_{2n-1}$  and  $z_{2n}$  are positive, the use of  $|\cdot|$  (three times) in (28) does not change the result. By analyzing all the roots  $t_n$ , the thresholds were found and listed in Table III, where the last column shows the relation between the threshold  $\beta_k$  and the roots  $t_n$  shown in (29)–(31).

For the listed thresholds in Table III,  $t_n$  in (29)–(31) is never real and negative, however  $t_n$  can be either complex or real with  $0 \leq t_n < 2$  for some  $\gamma < \gamma_0$ , resulting in virtual thresholds  $\beta_k$  and  $\beta_{8-k}$ . In what follows, we show that these thresholds are equal to each other when using (28), i.e., they fulfill the conditions in Theorem 2. First, consider the case when  $t_n$  is real but  $0 \leq t_n < 2$ . In this case  $z_{2n-1}$  and  $z_{2n}$  are complex with unit magnitude and according to (27), the corresponding thresholds  $\beta_k = -\beta_{8-k}$  are equal to zero. By analyzing the thresholds for all the RE and ARE patterns, we find that  $\beta_5 = -\beta_3$  for  $q = 3, 6, 9, 11$  and  $\beta_6 = -\beta_2$  for  $q = 5$ , which are separated by either no threshold or by the threshold  $\beta_4 = 0$ . These thresholds can be used in (9) according to Theorem 2.

Second, consider the case when  $t_n$  is complex. One of the other two roots of (48)  $t_{n'}$  giving  $\beta_{k'} = -\beta_{8-k'}$  is such that  $t_{n'} = t_n^*$ , which means that  $|t_n| = |t_{n'}|$ . This leads to two pairs of the thresholds  $\beta_k = \beta_{k'}$  and  $\beta_{8-k} = \beta_{8-k'}$ . Revising the thresholds for all the RE and ARE patterns we conclude that corresponding thresholds are:  $\beta_6 = -\beta_2$  and  $\beta_7 = -\beta_1$  for  $q = 8, 11$  and  $\beta_5 = -\beta_3$  and  $\beta_6 = -\beta_2$  for  $q = 7, 10$ . These thresholds can be used in (9) according to Theorem 2.

## REFERENCES

- [1] E. Zehavi, “8-PSK trellis codes for a Rayleigh channel,” *IEEE Trans. Commun.*, vol. 40, no. 3, pp. 927–946, May 1992.
- [2] G. Caire, G. Taricco, and E. Biglieri, “Bit-interleaved coded modulation,” *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.
- [3] A. Guillén i Fàbregas, A. Martinez, and G. Caire, “Bit-interleaved coded modulation,” *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 1–2, pp. 1–153, 2008.
- [4] M. K. Simon and R. Annavajjala, “On the optimality of bit detection of certain digital modulations,” *IEEE Trans. Commun.*, vol. 53, no. 2, pp. 299–307, Feb. 1988.
- [5] F. Gray, “Pulse code communications,” U. S. Patent 2632058, Mar. 1953.
- [6] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, “On the optimality of the binary reflected Gray code,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3170–3182, Dec. 2004.
- [7] E. Agrell, J. Lassing, E. G. Ström, and T. Ottosson, “Gray coding for multilevel constellations in Gaussian noise,” *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 224–235, Jan. 2007.
- [8] A. J. Viterbi, “An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes,” *IEEE J. Sel. Areas Commun.*, vol. 16, no. 2, pp. 260–264, Feb. 1998.
- [9] Ericsson, Motorola, and Nokia, “Link evaluation methods for high speed downlink packet access (HSDPA),” TSG-RAN Working Group 1 Meeting #15, TSGR1#15(00)1093, Tech. Rep., Aug. 2000.
- [10] P. Fertl, J. Jaldén, and G. Matz, “Performance assessment of MIMO-BICM demodulators based on mutual information,” *IEEE Trans. Sig. Proc.*, vol. 60, no. 3, pp. 1366–1382, Mar. 2012.
- [11] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2000.
- [12] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*. Prentice Hall, 1995.
- [13] K. Cho and D. Yoon, “On the general BER expression of one- and two-dimensional amplitude modulations,” *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, July 2002.
- [14] P. J. Lee, “Computation of the bit error rate of coherent  $M$ -ary PSK with Gray code bit mapping,” *IEEE Trans. Commun.*, vol. COM-34, no. 5, pp. 488–491, May 1986.
- [15] J. Lassing, E. G. Ström, E. Agrell, and T. Ottosson, “Computation of the exact bit-error rate of coherent  $M$ -ary PSK with Gray code bit mapping,” *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1758–1760, Nov. 2003.
- [16] J. Lassing, E. G. Ström, E. Agrell, and T. Ottosson, “Unequal bit-error protection in coherent  $M$ -ary PSK,” in *IEEE Vehicular Technology Conference (VTC-Fall)*, Orlando, FL, Oct. 2003.
- [17] L. Szczecinski, C. Gonzalez, and S. Aissa, “Exact expression for the BER of rectangular QAM with arbitrary constellation mapping,” *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 389–392, Mar. 2006.
- [18] J. Li, X. Zhang, Q. Gao, Y. Luo, and D. Gu, “Exact BEP analysis for coherent  $m$ -ary PAM and QAM over AWGN and Rayleigh fading channel,” in *IEEE Vehicular Technology Conference (VTC-Spring)*, Singapore, May 2008.
- [19] E. Agrell and A. Alvarado, “Optimal alphabets and binary labelings for BICM at low SNR,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6650–6672, Oct. 2011.
- [20] S. ten Brink, J. Speidel, and R.-H. Yan, “Iterative demapping and decoding for multilevel modulation,” in *IEEE Global Telecommunications Conference (GLOBECOM)*, Sydney, Australia, Nov. 1998.
- [21] P. Robertson, E. Villebrun, and P. Hoeher, “A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain,” in *IEEE International Conference on Communications (ICC)*, June 1995.
- [22] K. Hyun and D. Yoon, “Bit metric generation for Gray coded QAM signals,” *IEEE Proc.-Commun.*, vol. 152, no. 6, pp. 1134–1138, Dec. 2005.
- [23] M. S. Raju, R. Annavajjala, and A. Chockalingam, “BER analysis of QAM on fading channels with transmit diversity,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 481–486, Mar. 2006.
- [24] M. Benjillali, L. Szczecinski, S. Aissa, and C. Gonzalez, “Evaluation of bit error rate for packet combining with constellation rearrangement,” *Wiley Journal Wireless Comm. and Mob. Comput.*, pp. 831–844, Sep. 2008.
- [25] C. Stierstorfer, “A bit-level-based approach to coded multicarrier transmission,” Ph.D. dissertation, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, 2009.